

RAW SEQUENCE LISTING
PATENT APPLICATION US/09/421,213B

1642

#7

DATE: 04/07/2000
TIME: 12:04:31

Input Set: I421213B.RAW

This Raw Listing contains the General Information
Section and up to first 5 pages.

1 <110> APPLICANT: O'Brien, Timothy J.
2 <120> TITLE OF INVENTION: TADG-15: An Extracellular Serine Protease
3 Overexpressed in Carcinomas
4 <130> FILE REFERENCE: D6064CIP
5 <140> CURRENT APPLICATION NUMBER: US/09/421,213B
6 <141> CURRENT FILING DATE: 1998-10-20
7 <150> EARLIER APPLICATION NUMBER: US 09/027,337
8 <151> EARLIER FILING DATE: 02-20-1998
9 <160> NUMBER OF SEQ ID NOS: 98
10 <170> SOFTWARE: MS Word 98
11 <210> SEQ ID NO 1
12 <211> LENGTH: 3147
13 <212> TYPE: DNA
14 <213> ORGANISM: Homo sapiens
15 <220> FEATURE:
16 <223> OTHER INFORMATION: TADG-15
17 <400> SEQUENCE: 1
18 tcaagagcgg cctcggggta ccatggggag cgatcggggc cgcaagggcg gagggggccc 60
19 gaaggacttc ggcgcgggac tcaagtacaa ctcccggcac gagaaagtga atggcttgga 120
20 ggaaggcgtg gagttcctgc cagtcaacaa cgtcaagaag gtggaaaagc atggcccggg 180
21 gcgctgggtg gtgctggcag ccgtgctgat cggcctcctc ttggtcttgc tggggatcgg 240
22 cttcctgggtg tggcatttgc agtaccggga cgtgcgtgtc cagaaggtct tcaatggcta 300
23 catgaggatc acaaatgaga attttgtgga tgcctacgag aactccaact ccactgagtt 360
24 tgtaagcctg gccagcaagg tgaaggacgc gctgaagctg ctgtacagcg gagtcccatt 420
25 cctggggccc taccacaagg agtcggctgt gacggccttc agcgagggca gcgtcatcgc 480
26 ctactactgg tctgagttca gcatcccga gacactgggtg gaggaggccg agcgcgatcat 540
27 ggccgaggag cgcgtagtca tgctgcccc gcgggcgcgc tccctgaagt cttttgtggt 600
28 cacctcagtg gtggctttcc ccacggactc caaacagta cagaggaccc aggacaacag 660
29 ctgcagcttt ggctgcacg cccgcggtgt ggagctgatg cgcttcacca cgcccggctt 720
30 ccctgacagc ccctaccccg ctcattgccc ctgccagtg gcccgtgcgg gggacgccga 780
31 ctcagtgtcg agcctcacct tccgcagctt tgaccttgcg tccctgcgag agcgcggcag 840
32 cgacctgggt acggtgtaca acaccctgag ccccatggag cccacgccc tgggtgcagtt 900
33 gtgtggcacc taccctccct cctacaacct gaccttcac tccctccaga acgtcctgct 960
34 catcacactg ataaccaaca ctgagcggcg gcatcccggc tttagaggcca cttctttcca 1020
35 gctgcctagg atgagcagct gtggaggccg cttacgtaaa gccaggggga cattcaacag 1080
36 ccctactac ccaggccact accacccaa cattgactgc acatggaaca ttgagggtgcc 1140
37 caacaaccag catgtgaagg tgagcttcaa attctctac ctgctggagc ccggcgtgcc 1200
38 tgcgggcacc tgccccaagg actacgtgga gatcaatggg gagaaatact gcggagagag 1260
39 gtcccagttc gtgcgtacca gcaacagcaa caagatcaca gttcgcttcc actcagatca 1320
40 gtctacacc gacaccggct tcttagctga atacctctcc tacgactcca gtgacccatg 1380
41 cccggggcag ttcacgtgcc gcacggggcg gtgtatccgg aaggagctgc gctgtgatgg 1440
42 ctgggcccga tgcaccgacc acagcgatga gctcaactgc agttgcgagc ccggccacca 1500
43 gttcacgtgc aagaacaagt tctgcaagcc cctcttctgg gtctgcgaca gtgtgaacga 1560
44 ctgcccggag aacagcgaag agcagggggtg cagttgtccg gccagacct tcagggtgtc 1620

ENTERED

PAGE: 2

RAW SEQUENCE LISTING PATENT APPLICATION US/09/421,213B

DATE: 04/07/2000
TIME: 12:04:31

Input Set: I421213B.RAW

```

45 caatgggaag tgcctctcga aaagccagca gtgcaatggg aaggacgact gtggggacgg 1680
46 gtccgacgag gcctcctgcc ccaaggtgaa cgctcgtcact tgtaccaaac acacctaccg 1740
47 ctgcctcaat gggctctgct tgagcaaggg caaccctgag tgtgacggga aggaggactg 1800
48 tagcgacggc tcagatgaga aggactgcga ctgtgggctg cggtcattca cgagacaggc 1860
49 tcgtgttgtt gggggcacgg atgcggatga gggcgagtgg ccctggcagg taagcctgca 1920
50 tgctctgggc cagggccaca tctgcggtgc ttccctcatc tctcccaact ggctgggtctc 1980
51 tgccgcacac tgctacatcg atgacagagg attcaggtag tcagacccca cgcagtggac 2040
52 ggcttctctg ggcttgcaag accagagcca gcgcagcgcc cctgggggtgc aggagcgag 2100
53 gctcaagcgc atcatctccc accccttctt caatgacttc accttcgact atgacatcgc 2160
54 gctgctggag ctggagaaac cggcagagta cagctccatg gtgcggccca tctgcctgcc 2220
55 ggacgcctcc catgtcttcc ctgccggcaa ggccatctgg gtcacgggct ggggacacac 2280
56 ccagtatgga ggcaactggc cgctgatcct gcaaaaagggt gagatccgcg tcatcaacca 2340
57 gaccacctgc gagaacctcc tgccgcagca gatcacgccc cgcagtatgt gcgtgggctt 2400
58 cctcagcggc ggctgtggact cctgccaggg tgattccggg ggacccctgt ccagcgtgga 2460
59 ggcggatggg cggatcttcc aggcgggtgt ggtgagctgg ggagacggct gcgctcagag 2520
60 gaacaagcca ggcgtgtaca caaggctccc tctgtttcgg gactggatca aagagaacac 2580
61 tggggtatag gggccggggc caccctaatg tgtacacctg cggggccacc catcgtccac 2640
62 cccagtgtgc acgcctgcag gctggagact ggaccgctga ctgcaccagc gccccagaa 2700
63 catacactgt gaactcaatc tccagggtc caaatctgcc tagaaaacct ctcgttctct 2760
64 cagcctccaa agtggagctg ggaggtagaa ggggaggaca ctggtggttc tactgacca 2820
65 actgggggca aaggtttgaa gacacagcct cccccgccag ccccaagctg ggccgaggcg 2880
66 cgtttgtgta tatctgcctc ccctgtctgt aaggagcagc gggaacggag cttcggagcc 2940
67 tcctcagtga agtggtggg gctgccgat ctgggctgtg gggcccttgg gccacgctct 3000
68 tgaggaagcc caggctcgga ggacccctgga aaacagacgg gtctgagact gaaattgttt 3060
69 taccagctcc cagggtggac ttcagtgtgt gtatttgtgt aaatgggtaa aacaatttat 3120
70 ttctttttta aaaaaaaaaa aaaaaaa 3147

```

71 <210> SEQ ID NO 2

72 <211> LENGTH: 855

73 <212> TYPE: PRT

74 <213> ORGANISM: Homo sapiens

75 <220> FEATURE:

76 <223> OTHER INFORMATION: TADG-15

77 <400> SEQUENCE: 2

```

78 Met Gly Ser Asp Arg Ala Arg Lys Gly Gly Gly Gly Pro Lys Asp
79           5              10              15
80 Phe Gly Ala Gly Leu Lys Tyr Asn Ser Arg His Glu Lys Val Asn
81           20              25              30
82 Gly Leu Glu Glu Gly Val Glu Phe Leu Pro Val Asn Asn Val Lys
83           35              40              45
84 Lys Val Glu Lys His Gly Pro Gly Arg Trp Val Val Leu Ala Ala
85           50              55              60
86 Val Leu Ile Gly Leu Leu Leu Val Leu Leu Gly Ile Gly Phe Leu
87           65              70              75
88 Val Trp His Leu Gln Tyr Arg Asp Val Arg Val Gln Lys Val Phe
89           80              85              90
90 Asn Gly Tyr Met Arg Ile Thr Asn Glu Asn Phe Val Asp Ala Tyr
91           95             100             105
92 Glu Asn Ser Asn Ser Thr Glu Phe Val Ser Leu Ala Ser Lys Val
93          110             115             120
94 Lys Asp Ala Leu Lys Leu Leu Tyr Ser Gly Val Pro Phe Leu Gly

```

PAGE: 3

RAW SEQUENCE LISTING PATENT APPLICATION US/09/421,213B

DATE: 04/07/2000
TIME: 12:04:31

Input Set: I421213B.RAW

95		125		130		135
96	Pro Tyr His Lys Glu Ser Ala Val Thr Ala Phe Ser Glu Gly Ser					
97		140		145		150
98	Val Ile Ala Tyr Tyr Trp Ser Glu Phe Ser Ile Pro Gln His Leu					
99		155		160		165
100	Val Glu Glu Ala Glu Arg Val Met Ala Glu Glu Arg Val Val Met					
101		170		175		180
102	Leu Pro Pro Arg Ala Arg Ser Leu Lys Ser Phe Val Val Thr Ser					
103		185		190		195
104	Val Val Ala Phe Pro Thr Asp Ser Lys Thr Val Gln Arg Thr Gln					
105		200		205		210
106	Asp Asn Ser Cys Ser Phe Gly Leu His Ala Arg Gly Val Glu Leu					
107		215		220		225
108	Met Arg Phe Thr Thr Pro Gly Phe Pro Asp Ser Pro Tyr Pro Ala					
109		230		235		240
110	His Ala Arg Cys Gln Trp Ala Leu Arg Gly Asp Ala Asp Ser Val					
111		245		250		255
112	Leu Ser Leu Thr Phe Arg Ser Phe Asp Leu Ala Ser Cys Asp Glu					
113		260		265		270
114	Arg Gly Ser Asp Leu Val Thr Val Tyr Asn Thr Leu Ser Pro Met					
115		275		280		285
116	Glu Pro His Ala Leu Val Gln Leu Cys Gly Thr Tyr Pro Pro Ser					
117		290		295		300
118	Tyr Asn Leu Thr Phe His Ser Ser Gln Asn Val Leu Leu Ile Thr					
119		305		310		315
120	Leu Ile Thr Asn Thr Glu Arg Arg His Pro Gly Phe Glu Ala Thr					
121		320		325		330
122	Phe Phe Gln Leu Pro Arg Met Ser Ser Cys Gly Gly Arg Leu Arg					
123		335		340		345
124	Lys Ala Gln Gly Thr Phe Asn Ser Pro Tyr Tyr Pro Gly His Tyr					
125		350		355		360
126	Pro Pro Asn Ile Asp Cys Thr Trp Asn Ile Glu Val Pro Asn Asn					
127		365		370		375
128	Gln His Val Lys Val Ser Phe Lys Phe Phe Tyr Leu Leu Glu Pro					
129		380		385		390
130	Gly Val Pro Ala Gly Thr Cys Pro Lys Asp Tyr Val Glu Ile Asn					
131		395		400		405
132	Gly Glu Lys Tyr Cys Gly Glu Arg Ser Gln Phe Val Val Thr Ser					
133		410		415		420
134	Asn Ser Asn Lys Ile Thr Val Arg Phe His Ser Asp Gln Ser Tyr					
135		425		430		435
136	Thr Asp Thr Gly Phe Leu Ala Glu Tyr Leu Ser Tyr Asp Ser Ser					
137		440		445		450
138	Asp Pro Cys Pro Gly Gln Phe Thr Cys Arg Thr Gly Arg Cys Ile					
139		455		460		465
140	Arg Lys Glu Leu Arg Cys Asp Gly Trp Ala Asp Cys Thr Asp His					
141		470		475		480
142	Ser Asp Glu Leu Asn Cys Ser Cys Asp Ala Gly His Gln Phe Thr					
143		485		490		495
144	Cys Lys Asn Lys Phe Cys Lys Pro Leu Phe Trp Val Cys Asp Ser					

PAGE: 4

RAW SEQUENCE LISTING
PATENT APPLICATION US/09/421,213B

DATE: 04/07/2000
TIME: 12:04:31

Input Set: I421213B.RAW

145		500		505		510
146	Val Asn Asp Cys Gly Asp Asn Ser Asp Glu Gln Gly Cys Ser Cys					
147		515		520		525
148	Pro Ala Gln Thr Phe Arg Cys Ser Asn Gly Lys Cys Leu Ser Lys					
149		530		535		540
150	Ser Gln Gln Cys Asn Gly Lys Asp Asp Cys Gly Asp Gly Ser Asp					
151		545		550		555
152	Glu Ala Ser Cys Pro Lys Val Asn Val Val Thr Cys Thr Lys His					
153		560		565		570
154	Thr Tyr Arg Cys Leu Asn Gly Leu Cys Leu Ser Lys Gly Asn Pro					
155		575		580		585
156	Glu Cys Asp Gly Lys Glu Asp Cys Ser Asp Gly Ser Asp Glu Lys					
157		590		595		600
158	Asp Cys Asp Cys Gly Leu Arg Ser Phe Thr Arg Gln Ala Arg Val					
159		605		610		615
160	Val Gly Gly Thr Asp Ala Asp Glu Gly Glu Trp Pro Trp Gln Val					
161		620		625		630
162	Ser Leu His Ala Leu Gly Gln Gly His Ile Cys Gly Ala Ser Leu					
163		635		640		645
164	Ile Ser Pro Asn Trp Leu Val Ser Ala Ala His Cys Tyr Ile Asp					
165		650		655		660
166	Asp Arg Gly Phe Arg Tyr Ser Asp Pro Thr Gln Trp Thr Ala Phe					
167		665		670		675
168	Leu Gly Leu His Asp Gln Ser Gln Arg Ser Ala Pro Gly Val Gln					
169		680		685		690
170	Glu Arg Arg Leu Lys Arg Ile Ile Ser His Pro Phe Phe Asn Asp					
171		695		700		705
172	Phe Thr Phe Asp Tyr Asp Ile Ala Leu Leu Glu Leu Glu Lys Pro					
173		710		715		720
174	Ala Glu Tyr Ser Ser Met Val Arg Pro Ile Cys Leu Pro Asp Ala					
175		725		730		735
176	Ser His Val Phe Pro Ala Gly Lys Ala Ile Trp Val Thr Gly Trp					
177		740		745		750
178	Gly His Thr Gln Tyr Gly Gly Thr Gly Ala Leu Ile Leu Gln Lys					
179		755		760		765
180	Gly Glu Ile Arg Val Ile Asn Gln Thr Thr Cys Glu Asn Leu Leu					
181		770		775		780
182	Pro Gln Gln Ile Thr Pro Arg Met Met Cys Val Gly Phe Leu Ser					
183		785		790		795
184	Gly Gly Val Asp Ser Cys Gln Gly Asp Ser Gly Gly Pro Leu Ser					
185		800		805		810
186	Ser Val Glu Ala Asp Gly Arg Ile Phe Gln Ala Gly Val Val Ser					
187		815		820		825
188	Trp Gly Asp Gly Cys Ala Gln Arg Asn Lys Pro Gly Val Tyr Thr					
189		830		835		840
190	Arg Leu Pro Leu Phe Arg Asp Trp Ile Lys Glu Asn Thr Gly Val					
191		845		850		855

192 <210> SEQ ID NO 3
193 <211> LENGTH: 256
194 <212> TYPE: PRT

PAGE: 5

RAW SEQUENCE LISTING PATENT APPLICATION US/09/421,213B

DATE: 04/07/2000
TIME: 12:04:31

Input Set: I421213B.RAW

```

195 <213> ORGANISM: Homo sapiens
196 <220> FEATURE:
197 <223> OTHER INFORMATION: Hepsin
198 <400> SEQUENCE: 3
199   Arg Ile Val Gly Gly Arg Asp Thr Ser Leu Gly Arg Trp Pro Trp
200           5              10              15
201   Gln Val Ser Leu Arg Tyr Asp Gly Ala His Leu Cys Gly Gly Ser
202           20              25              30
203   Leu Leu Ser Gly Asp Trp Val Leu Thr Ala Ala His Cys Phe Pro
204           35              40              45
205   Glu Arg Asn Arg Val Leu Ser Arg Trp Arg Val Phe Ala Gly Ala
206           50              55              60
207   Val Ala Gln Ala Ser Pro His Gly Leu Gln Leu Gly Val Gln Ala
208           65              70              75
209   Val Val Tyr His Gly Gly Tyr Leu Pro Phe Arg Asp Pro Asn Ser
210           80              85              90
211   Glu Glu Asn Ser Asn Asp Ile Ala Leu Val His Leu Ser Ser Pro
212           95             100             105
213   Leu Pro Leu Thr Glu Tyr Ile Gln Pro Val Cys Leu Pro Ala Ala
214           110            115            120
215   Gly Gln Ala Leu Val Asp Gly Lys Ile Cys Thr Val Thr Gly Trp
216           125            130            135
217   Gly Asn Thr Gln Tyr Tyr Gly Gln Gln Ala Gly Val Leu Gln Glu
218           140            145            150
219   Ala Arg Val Pro Ile Ile Ser Asn Asp Val Cys Asn Gly Ala Asp
220           155            160            165
221   Phe Tyr Gly Asn Gln Ile Lys Pro Lys Met Phe Cys Ala Gly Tyr
222           170            175            180
223   Pro Glu Gly Gly Ile Asp Ala Cys Gln Gly Asp Ser Gly Gly Pro
224           185            190            195
225   Phe Val Cys Glu Asp Ser Ile Ser Arg Thr Pro Arg Trp Arg Leu
226           200            205            210
227   Cys Gly Ile Val Ser Trp Gly Thr Gly Cys Ala Leu Ala Gln Lys
228           215            220            225
229   Pro Gly Val Tyr Thr Lys Val Ser Asp Phe Arg Glu Trp Ile Phe
230           230            235            240
231   Gln Ala Ile Lys Thr His Ser Glu Ala Ser Gly Met Val Thr Gln
232           245            250            255
233   Leu
234 <210> SEQ ID NO 4
235 <211> LENGTH: 225
236 <212> TYPE: PRT
237 <213> ORGANISM: Homo sapiens
238 <220> FEATURE:
239 <223> OTHER INFORMATION: SCCE
240 <400> SEQUENCE: 4
241   Lys Ile Ile Asp Gly Ala Pro Cys Ala Arg Gly Ser His Pro Trp
242           5              10              15
243   Gln Val Ala Leu Leu Ser Gly Asn Gln Leu His Cys Gly Gly Val
244           20              25              30

```

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Input Set: I421213B.RAW

Line ? Error/Warning

Original Text

622 W "N" or "Xaa" used: Feature required

tgggtngtna cngcngcnca ytg

632 W "N" or "Xaa" used: Feature required

arnggnccnc cnswrtncc